

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 16, 2001, 18:14:46 ; Search time 45.43 Seconds  
(without alignments)  
1872.598 Million cell updates/sec

Title: US-09-405-504A-53  
Perfect score: 3384  
Sequence: 1 MLLGASLVGLLFSKLVKL.....RYVPLDQAYSIQAGEEKL 643

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- SPTREMBL\_16:\*
- 1: sp\_archaea:\*
  - 2: sp\_bacteria:\*
  - 3: sp\_fungi:\*
  - 4: sp\_human:\*
  - 5: sp\_invertebrate:\*
  - 6: sp\_mammal:\*
  - 7: sp\_mhc:\*
  - 8: sp\_organelle:\*
  - 9: sp\_phage:\*
  - 10: sp\_plant:\*
  - 11: sp\_rodent:\*
  - 12: sp\_unclassified:\*
  - 13: sp\_vertebrate:\*
  - 14: sp\_virus:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	3152	93.1	641	4	O95186
2	2433	71.9	506	11	O88562
3	1427.5	42.2	671	5	O9w1v9
4	1408.5	41.6	690	5	O9VKU1
5	1292.5	38.2	661	5	O9W185
6	1292	38.2	650	5	O19878
7	1281	37.9	655	5	O18916
8	1122	33.2	619	4	O9Y2P4
9	1053.5	31.1	614	11	O88561
10	1021	30.2	620	11	O70550
11	984.5	29.1	690	4	O9Y2P4
12	969.5	28.6	690	11	O9E338
13	960	28.4	662	11	O88694
14	918.5	27.1	608	2	O9HZV4
15	886.5	26.2	597	2	O05307
16	877.5	25.9	623	11	O88560
17	801	23.7	643	3	O42633
18	742	21.9	669	3	O60021
19	408	12.1	502	2	O53551

20	389	11.5	503	2	O53306	O53306 mycobacteri
21	375	11.1	569	1	O28502	O28502 archaeoglob
22	371.5	11.0	555	2	O85737	O85737 streptomyce
23	364.5	10.8	513	2	O07610	O07610 bacillus su
24	358	10.6	532	2	P72007	P72007 mycobacteri
25	346.5	10.2	535	2	O9KHL1	O9KHL1 streptomyce
26	332	9.8	599	1	O9Y9K4	O9Y9K4 aeropyrum p
27	322.5	9.5	560	2	P95227	P95227 mycobacteri
28	318.5	9.4	504	2	O9F7P5	O9F7P5 uncultured
29	318	9.4	661	3	O13441	O13441 coprinus ci
30	311.5	9.2	549	2	O31826	O31826 bacillus su
31	310	9.2	2723	2	O30479	O30479 streptomyce
32	301.5	8.9	486	2	O34837	O34837 bacillus su
33	300.5	8.9	549	2	O9R915	O9R915 bacillus su
34	298.5	8.8	593	1	O29233	O29233 archaeoglob
35	297	8.8	545	10	O9SWH8	O9SWH8 solanum tub
36	296.5	8.8	511	2	O9KZC1	O9KZC1 streptomyce
37	294.5	8.7	552	1	O28762	O28762 archaeoglob
38	294	8.7	542	10	O9FOY7	O9FOY7 capsicum an
39	294	8.7	554	2	O07411	O07411 mycobacteri
40	293.5	8.7	512	3	O74976	O74976 schizosacch
41	292	8.6	584	2	O9RTR4	O9RTR4 deinococcus
42	291	8.6	444	2	O9X4W6	O9X4W6 pseudomonas
43	291	8.5	535	10	O81139	O81139 populus tre
44	286	8.5	8563	2	O54297	O54297 streptomyce
45	285.5	8.4	517	2	O9JZ62	O9JZ62 neisseria m

## ALIGNMENTS

### RESULT 1

O95186	AC	O95186	PRELIMINARY; PRT; 641 AA.
DT	01-MAY-1999	(Tremblrel. 10, Created)	
DT	01-MAY-1999	(Tremblrel. 10, Last sequence update)	
DT	01-MAY-2000	(Tremblrel. 13, Last annotation update)	
DE	FATTY ACID TRANSPORT PROTEIN.		
GN	FATP.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=HEART;		
RX	MEDLINE=9096471; PubMed=9878842;		
RA	Fitscher B.A., Riedel H.D., Young K.C., Stremmel W.;		
RT	"Tissue distribution and cDNA cloning of a human fatty acid transport		
RL	protein (hsFATP4).";		
RL	Biochim. Biophys. Acta 1443:381-385(1998).		
DR	EMBL; AF055899; AAD11623.1; -		
DR	InterPro; IPR000566; -		
DR	InterPro; IPR000873; -		
DR	Pfam; PF00501; AMP-binding; 1.		
DR	PROSITE; PS00455; AMP_BINDING; 1.		
DR	PROSITE; PS00213; LIPOCALIN; UNKNOWN_1.		
SQ	SEQUENCE 641 AA; 71431 MW; 8C24F76C9BF81378 CRC64;		

Query Match 93.1%; Score 3152; DB 4; Length 641;  
Best Local Similarity 94.7%; Pred. No. 6.6e-256;  
Matches 610; Conservative 4; Mismatches 26; Indels 4; Gaps 3;

OY	1	MLLGASLVGLLFSKLVKLPTQVGFSLFLYLGSGWRFIRVFIKTRDIFGGLVLL 60	
Db	1	MLLGASLVGLLFSKLVKLPTQVGFSLFLYLGSGWRFIRVFIKTRDIFGGLVLL 60	
OY	61	KVAKVRQCLOQRTPVILPASTVRRHPDKTALIFGTDHTWTFRLDEYSSVANFLQA 120	
Db	61	KVAKVRQCLOQRTPVILPASTVRRHPDKTALIFGTDHTWTFRLDEYSSVANFLQA 120	

FT	NON_TER	1	1	71.9%;	Score 2433;	DB 11;	Length 506;
FT	CARBOHYD	190	N-LINKED (GLCNAC. . .)	(POTENTIAL).			
FT	CARBOHYD	253	N-LINKED (GLCNAC. . .)	(POTENTIAL).			
FT	CARBOHYD	378	N-LINKED (GLCNAC. . .)	(POTENTIAL).			
SEQ	SEQUENCE	506 AA;	56902 MR;	4E9BC80C2173326D	CRC64;		
Query Match							
Best Local Similarity 91.2%; Pred. No. 1.le-195;							
Matches 455; Conservative 23; Mismatches 21; Indels 0; Gaps							
QY	145	GNAKLGVPAALINTNLRRDALLHCLTTSRARALVFGSEMASAICEVHASLDPSLSLPCSG	204				
DB	8	GNAKLGVPAALINTNLRRDALLHCLTTSRARALVFGSEMASAICEVHASLEPTLSLPCSG	67				
QY	205	SNEPGAVPSTEHLDPLLKDAPKHLSPCKDGTDKLFYITSGTGLPKAAIVVHSRY	264				
DB	68	SNEPTVPSTEHLDPLLEDAPKHLSPCKDGTDKLFYITSGTGLPKAAIVVHSRY	127				
QY	265	RMAALVYVGFMRPNDIVDCLPLYSAGNIVIGQCLLHGMTVVRKKSASRFWDCCI	324				
DB	128	RMAALVYVGFMRPNDIVDCLPLYHSSRKHRRGDMQCLLHGMTVVRKKSASRFWDCCI	187				
QY	325	KYNCTIVQYIGELRYLLNPPREANQHVRMALGNGLRQSIWTFNSSRPHIPQVAEY	384				
DB	188	KYNCTVVOYIGELRYLLNPPREASRHKVRMALGNGLRQSIWTFSSRPHIPQVAEY	247				
QY	385	GATECNCSLGNFDSQVGCAGFNSRILSFVYPIRLVRVNEDTMELLRGPDGYCIPCQGP	444				
DB	248	GATECNCSLGNFDSRVGCAGFNSRILSFVYPIRLVRVNEDTMELLRGPDGYCIPCQGP	307				
QY	445	GOLVGRILIQDPLRRFDGYLNOGANNKIAKOVFKKGQOAYLTGDLVMDLGYLYFRDR	504				
DB	308	GOLVGRILIQDPLRRFDGYLNOGANNKIIANDVEFKGQOAYLTGDLVMDLGYLYFRDR	367				
QY	505	TGDTFRWKGENVSTTEVETGLSRLLDMADVNYGVGVPGTEGRAGMAVAVSPTGNCDLR	564				
DB	368	TGDTFRWKGENVSTTEVETGLSRLLHMADVNYGVGVPGTEGRAGMAVAVSPTGNCDLR	427				
QY	565	FAQVLEKELPLYPARIFLRLPELPHKTGTGVYKFOKTELKREGFDPDAIVKDPFLFYLDAAQGR	624				
DB	428	FAQTLKELPLYPARIFLRLPELPHKTGTGVYKFOKTELKREGFDPVSKDPFLFYLDARKGC	487				
QY	625	YVPLDQEAYSRIQAEEKL	643				
DB	488	YVALDQEAYTRIOAEEKL	506				

Db 368

Qy	565	FAQVLEKELPYARPIFLRLPELHKTGYTKFQKTELKEGDFDAIVKDPLFYDLDAOKGR	624
Dd	428	EAQTLLKELPYARPIFLRELPELHKTGTFFKQKTELKEGFDPSSVVAKDPLFYLDARKGC	487
Qy	625	YVPQLQEAVSRIOAGEEKL 643	
Dd	488	YVALDQAYTRIOAGEEKL 506	
RESULT	3		
Q9W1v9			
ID	Q9W1v9	PRELIMINARY;	PRT; 671 AA.
AC	Q9W1v9;		
DT	01-MAY-2000 (TREMBLrel. 13, Created)		
DT	01-NAR-2001 (TREMBLrel. 16, Last sequence update)		
DT	01-NAR-2001 (TREMBLrel. 16, Last annotation update)		
DE	CG3037 PROTEIN.		
DE	CG3037		
GN	Crosophila melanogaster (Fruit fly).		
OS	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;		
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;		
OC	Ephydroidea; Drosophilidae; Drosophila.		
OX	NCBI_TaxID=7227;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=BERKELEY;		
RX	MEDLINE=20196006; PubMed=10731132;		
RA	Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,		
RA	Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,		
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,		
RA	Sutton G.G., Wortman J.R., Yeung M.D., Zhang Q., Chen L.X.,		
RA	Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champé M., Pfeiffer B.D.,		
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,		
RA	Abriél J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,		

[illegible]

Db	435	ISRIPLQIYPISIIKADPHTGPELRNSQGLCEADPEGVFGKIVRGNPCREFLYVD	494
Qy	466	QGANKKIAKDVFKKQDQAYLTGDLVMDLGYLYIFDRDGTDFRNKGENVSTTEVEGTL	525
Db	495	QKASSKKVVHDFVSKGDMAFISGDLVADERGLYIFADRTGDTFRNKGENVSTSEYEQ	554
Qy	526	SLRLDMADVYGVVEVPGTEGRAGMAAVASPTGNCMLDRFAQVLEKELPLYARPIFLRL	585
Db	555	SNLAGYKQVIYVGSIPNTEGRAGMAAIYDPTREVNYSQGLVELAKSLPNYARQFLRL	614
Qy	586	PELHGTGTYKQKTELKKEGDFPAIVKDPFLFYLDQAQGRVPLDQAYSRIOAGE	640
Db	615	RKIDLTGFKURKVELQQQGNPEIIDELFYAQPD-GVYAPLTPSVYRIVRNE	668
RESULT	4		
Q9VKU1		PRELIMINARY;	690 AA.
AC	Q9VKU1		
DT	01-MAY-2000	(TrEMBLrel. 13, Created)	
DT	01-MAY-2000	(TrEMBLrel. 13, Last sequence update)	
DT	01-MAR-2001	(TrEMBLrel. 16, Last annotation update)	
DE	CG7400	PROTEIN.	
GN	FATP	OR CG7400.	
OS	Drosophila melanogaster	(Fruit fly).	
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;		
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;		
OC	Ephydroidea; Drosophilidae; Drosophila.		
OX	NCBI_TaxID:7227;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-BERKELEY;		
RX	MEDLINE=20196006; PubMed=10731132;		
RA	Adams M.D., Celnikier S.E., Holt R.A., Evans C.A., Gocayne J.D.,		
RA	Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,		
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,		
RA	Stouten G.G., Wortman J.R., Vandell M.D., Zhang Q., Chen L.X.,		
RA	Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champagne M., Pfeiffer B.D.,		
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,		
RA	Abriel J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,		
RA	Ballew R.M., Basu A., Bakendale J., Bayraktaroglu L., Beasley E.M.,		
RA	Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,		
RA	Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,		
RA	Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,		
RA	Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,		
RA	de Pablo S., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,		
RA	Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,		
RA	Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,		
RA	Fosler C., Gabrielsen A.E., Garg N.S., Gelbart W.M., Glasser K.,		
RA	Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,		
RA	Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,		
RA	Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,		
RA	Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,		
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,		
RA	Lasko P., Lei Y., Levitsky A.C., Li J., Li Z., Liang Y., Lin X.,		
RA	Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,		
RA	Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,		
RA	Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,		
RA	Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,		
RA	Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,		
RA	Reinartz K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,		
RA	Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,		
RA	Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,		
RA	Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,		
RA	Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,		
RA	Williams S.M., Woodgate T., Worley K.C., Wu D., Yang S., Yao Q.A.,		
RA	Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,		
RA	Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,		
RA	Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;		
RT	"The genome sequence of Drosophila melanogaster."		
RL	Science 287:2185-2195(2000).		
DR	EMBL; AE003629; AAF52969.1; -		
DR	FlyBase; FBgn0021953; Fatp.		



QY 386 ATECNCSLGNFDSQVAGCNFSRILSFVYPIRLVVRVNETMELIRGDCVPCOPGEPG 445  
 DB 405 ATEGNSNLINTRVGAIGFVYVYVQVLRDEVTGELLKDSKGCRCOPGQAG 464  
 QY 446 OLVGRIQKDLRRFGYLNQGNANKKIAKDFKGDQAYLGTDLVMDLGLYLFRRRT 505  
 DB 465 LVGVKVDARRAVAFHGYADKGASEQKLLRNVTSGDVFNSGDMVVRDILGYFFKORT 524  
 QY 506 GDFRKNKGENVSTVEGTLRLDMDADVAVYGVVEVPGTEGRAGMAVAPSTGNCDLERF 565  
 DB 525 GDFRKNKGENVSTVEGTLRLDMDADVAVYGVVEVPGTEGRAGMAVAPSTGNCDLERF 565  
 QY 566 AOVLEKELPIYARPIRLPLPHKGTGKFOKTELKKEGPDPAIVKDPFLFYLDQAQKGRY 625  
 DB 585 SVVLRSGLPPYARPIRLPLPHKGTGKFOKTELKKEGPDPAIVKDPFLFYLDQAQKGRY 625  
 QY 626 VPLDQAYSRIOAGEKL 643  
 DB 644 RPLSQOHELLRSKGAGL 661

RESULT 6  
 Q19878 ID Q19878 PRELIMINARY; PRT; 650 AA.  
 AC Q19878  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)  
 DE FATTY ACID TRANSPORTER PROTEIN A (F28D1.9 PROTEIN).  
 GN F28D1.9.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Baynes C.;  
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: INVOLVED IN TRANSLOCATION OF LONG-CHAIN FATTY ACIDS (BY  
 CC SIMILARITY).  
 CC -1- SIMILARITY: TO OTHER ENZYMES WHICH ACT VIA AN ATP-DEPENDENT  
 CC COVALENT BINDING OF AMP TO THEIR SUBSTRATE.  
 DR EMBL; 270684; CAA94602.1;  
 DR WormPep; F28D1.9; CE05749.  
 DR InterPro; IPR000873;  
 DR Pfam; PF00501; AMP-binding; 1.  
 DR PROSITE; PS00455; AMP-BINDING; 1.  
 KW Glycoprotein; Lipid transport; Transmembrane; Transport.  
 FT TRANSMEM 23 43 POTENTIAL.  
 FT TRANSMEM 137 157 POTENTIAL.  
 FT TRANSMEM 286 306 POTENTIAL.  
 FT CARBOHYD 100 100 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 518 518 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 650 AA; 72759 MW; F2D16649EAL3288A CRC64;

Query Match 38.2%; Score 1292; DB 5; Length 650;  
 Best Local Similarity 41.8%; Pred. No. 9.7e-100;  
 Matches 272; Conservative 123; Mismatches 228; Indels 28; Gaps 10;  
 QY 7 LVGVLLFSKLV--KLPTQVGSLLFYLGSGGWRFTIRVFIKTRIDFGGLVLLKVA 64  
 DB 6 LVTVMLLTAVIAQNLPGVILAGVILYITVWHGDFIYRSYLTNRLDGLTGLALIEVKI 65  
 QY 65 KYRQCLQRRRTVPILFASTVRRHPDKTALIFEGDTHWTFQQLDEYSSVANFQARGLA 124  
 DB 66 DLWRLHQNKGHLEFLDIVKKNPNKPAIDTETNYAEFNAHCNRVANYFQGLGRY 125  
 QY 125 SGDVAIFMENRNEFFVGLWGLMAKLGVEAALINTNLRDALLHCLTTSRARALVFGSEMA 184  
 DB 126 SGDVALYMWENSVEVAWGLAKIGVVTAVINSNLKREQLVHCITASKYKAITSTVQLQ 185

QY 185 SAICEVHASLDPSL-----SLFCSGSWEP---GAVPSTPEHLDPKADPKHLPSPCDK 235  
 DB 186 NIMLD---AIDQKLFDEVEGIEVSVG--EPKNSGFKMLKKKLDQAITTEPKTLDIV--- 237  
 QY 236 GFTDKLFYITSGTGLPKAAIVVHSRYRMAALVYVYGFMRPNDIVYDCLPLVHSAGNI 295  
 DB 238 DFKSLICFYITSGTGMKFAAVMKHFYISYAVGAASFGIRPSDRMVYSMPYIHTAAGI 297  
 QY 296 VIGOCCLHGMVTVIRKFSASREFWDDCIKYNCTIVYIGELCYLLNQPPREAEHQV 355  
 DB 298 LGVGALLGGSCVIRKKFSASNEWDCVKYDCTVSYIGEICRYLLAQPWVEESHRM 357  
 QY 356 RMALGNLGRQSIWTFNFSRPHIPOVAFYGATECNSLGNPDSQVACGF--NSRILSFV 413  
 DB 358 RLLVGNGLRAEIQWPFVDRFRV--RIGELYGSTESTSLVNIHDHVGACGFLPSPLTKM 416  
 QY 414 YPIRLVRNEDTMEIRGDCVPCOPGEPGOLVGRIOKDLPLRFDGVLNQCANNKI 473  
 DB 417 HPVRLIKVDDVTGEAIRTSGLCIACAPGSGAMVSTRKNNPLLQFEGYLNKKTNNKI 476  
 QY 474 AKDVKKGDQAYLGTDLVMDLGLYFDRDTGDTFRWKGENVSTTEVEGTLRLDMD 533  
 DB 477 IRDVFAGKDCSFLGTDLHWDRLGVYVFKORTGDTFRWKGENVSTTEVEAILHPITGLSD 536  
 QY 534 VAVYGVVEVPGTEGRAGMAVAPSTGNC-----LERFAQVLEKELPIYARPIFLRLPEL 588  
 DB 537 ATVYGVVEVPGTEGRAGMAVAPSTGNC-----LERFAQVLEKELPIYARPIFLRLPEL 588  
 QY 589 HKTGYFKQTELKKEGFDPAIVKDPFLFYLDQAQGRVYVPLDQAYSRIOAG 639  
 DB 597 EKTGTFKLKTNLQRLGIMDA-PSDSIYIYNSNRNEVFPFNDLRCKVSLG 646

RESULT 7  
 Q18916 ID Q18916 PRELIMINARY; PRT; 655 AA.  
 AC Q18916  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)  
 DE FATTY ACID TRANSPORTER PROTEIN B.  
 GN DI009.1.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Favello T.;  
 RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Waterston R.;  
 RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: INVOLVED IN TRANSLOCATION OF LONG-CHAIN FATTY ACIDS (BY  
 CC SIMILARITY).  
 CC -1- SIMILARITY: TO OTHER ENZYMES WHICH ACT VIA AN ATP-DEPENDENT  
 CC COVALENT BINDING OF AMP TO THEIR SUBSTRATE.  
 DR EMBL; U40938; AAA81698.1;  
 DR HSSP; P14687; IAMU.  
 DR WormPep; DI009.1; CE04285.  
 DR InterPro; IPR000873;  
 DR Pfam; PF00501; AMP-binding; 1.  
 DR PROSITE; PS00455; AMP-BINDING; 1.  
 KW Glycoprotein; Lipid transport; Transmembrane; Transport.  
 FT TRANSMEM 10 30 POTENTIAL.  
 FT TRANSMEM 142 162 POTENTIAL.  
 FT TRANSMEM 295 315 POTENTIAL.  
 FT CARBOHYD 176 176 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 282 282 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 332 332 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 522 522 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 655 AA; 73514 MW; 2AEB3BF7B198D14C CRC64;

```
Query Match          37.9%; Score 1281; DB 5; Length 655;
Best Local Similarity 41.4%; Pred. No. 8.2e-99;
Matches 266; Conservative 116; Mismatches 230; Indels 30; Gaps 8;

QY 23 TQVGFSLFLYLGLSGGWRFT-----RVFIKTRRRDIFGGLVLKAKVR 67
DB 14 TEVYAVLVNVSVEWKFVIGYVVVRLRTDGRALATLPDRFAGLKLISYKSTIR 73

QY 68 QCLQERTVILFASIVRRHPDKTALIFEGTDTHWTFRLQDEYSSSVANFLQAG 127
DB 74 GLFKKDKPIHEILNQVQKPNVAIETESGRLTYQELNALANQYALNYSEGYKMG 133

QY 128 VAAIFMENRNEFVGLWGLMAKLGVEAALINTNRRDALLHCLTTSRARALVFG 184
DB 134 VVALFENSIDFAIWLGLSKIGVSAFINSNKLPLELAHSINVSCKCKIINILLPM 193

QY 185 SATCEVHASLDPSLSFCSSWPGAVPSTEHLDPLLDKAPKHLPCDPKGTOKLFYI 244
DB 194 KAAREKNLISD-EIHVFLAGTQVDRHSRSLQDLHLFSEDEP---EVIDGLNFRS 249

QY 245 YTSCTTGLPRAAIVVHSRYRMAALVYVGRMPNDIVYDCLPLYSAGNIVIGOCLLH 304
DB 250 YTSCTTGNPAPVIAKHFRYFWIANGAKAPGINKSDVYITMFWYHSAAGIWIGSL 309

QY 305 GMTVVIRKFSASREWDDCIYKNTCTIVQYIGELCYRLLNQPPREAEHQVHMALGN 364
DB 310 GSTAVIRKFSASRNWKCQVKNVTATQYIGETCYRLLAANPCPEKQHNVRMLWGN 369

QY 365 QSIWTFNSSRFHPIQVAFEGATECNSLGNFDSQVAGCGFNSRI--LSFVPIRL 422
DB 370 GQIWKEVGFGRGKIGELYGSTEGNSIVNVNHNHVGACGFMPYIPIHISLSPVRL 429

QY 423 EDTMELIRGPGVCIPQCPGEPQGLYGRITQKDLRRFDGLYLNQAGNKKIADVF 482
DB 430 RATGELERDNGLCVPCVPGCTGEMGVVKEIKDILLKFGYVSEGTAKKIYRDVFK 489

QY 483 QAYLTGVLVMDLGYLYFYRDRGDTFRWKGENSVTEVEGTLRLLLMDADVAVG 542
DB 490 KVFASGDIHLWDDLYLYFYFDRGDTFRWKGENSVTEVEGTLQPVVEDATVYGV 549

QY 543 GTEGRAGMAAVASTGNCLELRE---AOVLEKELPLYARPIRLPLHKLTKGTG 598
DB 550 KMEGRAGMAGIVVDGT-DVEKFIADTISRLTENLASYAIPVIRLCKEVDRTGIF 608

QY 599 TELRKEGDFPAIVK-DPLFLYDLAQGRYVPLDQAYSRIQAG 639
DB 609 TDLQKQGYDLVACKGDPFIYVWSAAEKSYKPLTDKMQQDIDTG 650

RESULT 8
QY2P4 PRELIMINARY; PRT; 619 AA.
AC QY2P4;
DT 01-NOV-1999 (TREMREL. 12, Created)
DT 01-NOV-1999 (TREMREL. 12, Last sequence update)
DT 01-MAR-2001 (TREMREL. 16, Last annotation update)
DE VERY LONG-CHAIN ACYL-COA SYNTHETASE HOMOLOG 1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Steinberg S.J., Watkins P.A.;
RT "Human Very Long-Chain Acyl-CoA Synthetase Homolog 1.";
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF064254; AD29443.1;
DR InterPro: IPR000873;
DR Pfam: PF00501; AMP-binding; 1.
DR PROSITE: PS00455; AMP_BINDING; 1.
SQ SEQUENCE 619 AA; 70111 MW; 20264CE3FBB44FF5 CRC64;

Query Match          33.2%; Score 1122; DB 4; Length 619;
Best Local Similarity 39.2%; Pred. No. 1.7e-85;
Matches 258; Conservative 107; Mismatches 230; Indels 64; Gaps 12;

QY 2 LIGASLVGLFLSLKVLKLPWTVQVGSLLFLYLGSGGWRFTIRVFIKTRRRDIFG 61
DB 8 VLGAGW-VLHFLQ-----KLLPFYFWDDEFVLKV-----VLII 41

QY 62 VKAKVRCQLQERRIVPILFASTVRRHPDKTALIFEGTDTHWTFRLQDEYSSSVAN-FLQA 120
DB 42 IRLKYEKRGELVTVLQKFLSHAKRQPKPKFIYIEG--DIYIQDVDRKSRVAHVLFNH 99

QY 121 RGLASGVAAIFMENRNEFVGLWGLMAKLGVEAALINTNRRDALLHCLTTSRARALV 180
DB 100 SSLKGGDITVALLMSNEPDEFVHWFGLAKLGCVAFIATNIRNSNLLNCIRACGPRAL 159

QY 181 SEMASACEVHASLDPSLSFCSG-----SWEPGAVPSTEHLDPLLDKAP 226
DB 160 ADLLGTVEEILFSLSENISVWGMKDSVPQGVISLKEKLSSTSPDFVPRSHHVSVLLKS 217

QY 227 KHLPCSPDKGFTDKLFYITTSCTTGLPKAAIVVHSRYRMAALVYVGRMPNDIVYDCL 286
DB 218 ----TC-----LYIFTSGTGLPKAAVISQIQLVRGSA-VLWAFGCTAHDIVYITL 263

QY 287 PLYHSAGNIVIGOCLLHGMTVIRKFSASREWDDCIYKNTCTIVQYIGELCYRLLNQPP 346
DB 264 PLYHSSAAILGISGCVELGATCVLKKKFSASQFWSQDKCKYDVTVFQYIGELCYR 323

QY 347 REAENQHOVRMALGNLGRQSIWTFNSSRFHPIQVAFEGATECNSLGNFDSQVAGCG 406
DB 324 REGEKDHVRALIGNHRSDVWREFLGRFNGIKVCELYAATESISFMNVTGRIGART 383

QY 407 SRILSFVYPIRLVRVNEDELMELIRGPGVCIPQCPGEPQGLYGRITQKDLRRFDGL 466
DB 384 NLFYKLLSTFDLIKYDFOKDEPMRNEQGCWCIHVKKGEPGLLISRVNAKNP---PFG 440

QY 467 GANNK-KIADVFKKGGQAYLTGVLVMDLGYLYFYRDRGDTFRWKGENSVTEVEGTL 525
DB 441 YKHTKDKLLCDLVFKKGGVYLTGLTIVQDQDNFLYFWDRTGDTFRWKGENVATTEVA 500

QY 526 SRLIDMAADVAVGVPEVPCTEGRAGMAA-VASPTGNCLELREFAOVLKELPLYARPI 584
DB 501 GMLDPIQANVYGAISGYEGRAGMASILKPNLSLDERKVEYQVVTFLPAYACPRLRI 560

QY 585 LPELHKTKTYKFKTELKKEGDFPAIVKDPFLYDLAQGRYVPLDQAYSRIQAGEKL 643
DB 561 QEKMEATGCTKLLKHQLVEDGFNPLKISEPLYFMDNLKKSYYLLTRELIDQIMLGEIKL 619

RESULT 9
O88561 PRELIMINARY; PRT; 614 AA.
AC O88561;
DT 01-NOV-1998 (TREMREL. 08, Created)
DT 01-NOV-1998 (TREMREL. 08, Last sequence update)
DT 01-MAR-2001 (TREMREL. 16, Last annotation update)
DE FATTY ACID TRANSPORT PROTEIN 3 (FATP3) (LONG-CHAIN FATTY ACID
DE TRANSPORT PROTEIN 3) (FRAGMENT).
GN SIC27A3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=98337965; PubMed=9671728;
RA Hirsch D., Stahl A., Lodish H.F.;
RT "A family of fatty acid transporters conserved from mycobacterium to
RT man.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:8625-8629(1998).
CC -1- FUNCTION: INVOLVED IN TRANSLOCATION OF LONG-CHAIN FATTY ACIDS
```



CC ACROSS THE PLASMA MEMBRANE. MAY PLAY A PIVOTAL ROLE IN REGULATING  
CC AVAILABLE LONG-CHAIN FATTY ACID SUBSTRATES FROM EXOGENOUS SOURCES  
CC IN TISSUES UNDERGOING HIGH LEVELS OF BETA-OXIDATION OR  
CC TRIGLYCERIDE SYNTHESIS.  
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. PLASMA MEMBRANE.  
CC -1- TISSUE SPECIFICITY: LUNG, LIVER, AND TESTIS.  
CC -1- SIMILARITY: TO OTHER ENZYMES WHICH ACT VIA AN ATP-DEPENDENT  
CC COVALENT BINDING OF AMP TO THEIR SUBSTRATE.  
DR EMBL: AF072758; AAC40187.1; -  
DR MGD; MGI:1347358; SLC27a3.  
DR InterPro: IPR000873; -  
DR PRINTS; PF00501; AMP-binding; 1.  
DR PROSITE; PS00154; AMPBINDING.  
DR PROSITE; PS00455; AMP BINDING; 1.  
KW Glycoprotein; Lipid transport; Transmembrane; Transport.  
FT NON\_TER 1  
FT TRANSMEM 99 119 POTENTIAL.  
FT TRANSMEM 262 282 POTENTIAL.  
FT CARBOHYD 367 367 N-LINKED (GLCNAC... ) (POTENTIAL).  
SQ SEQUENCE 614 AA; 67041 MW; 33C2A558CDD9D989 CRC64;

Query Match 31.1%; Score 1053.5; DB 11; Length 614;  
Best Local Similarity 39.9%; Pred. No. 9.4e-80;  
Matches 244; Conservative 85; Mismatches 219; Indels 63; Gaps 11;

QY 81 ASTVRRHPDKTALIFEGTDHTWFRQLDEYSSVAN-FIQARG----- 122  
DB 19 AYLARQPHTHFLHGA--QRSYFAERESNRIARAFRLARGWTGRRSGRGSTEBGA 76  
QY 123 -----LASGDVAAFMFENRNEFVGLWGMALGLVEAAALINTNLREDA 164  
DB 77 RVAPPAGDAAARTGTTAPPLAGCATVALLPAGDFLWTFGLAKAGLTATFVPTALRGP 136  
QY 165 LLHCLTTSARALVFGSEMAISCEVHASLDPSLSFCGSGWEPGVPST- 215  
DB 137 LLHCLSCGASALVATEFLE-----SLEPDLPALRAMGLHLWATPETNAGISNLL 189  
QY 216 -EHLDPDLLDAPKLPSCPKGFTKLFVITVSGTGLPKAAIVVHSRYRMAALVYGF 274  
DB 190 SEAAQDVEPVGYL-SAP-QNIMDTCLYIFTSGITGLPKAARISHLKLVCQCGF-YHLC 246  
QY 275 RMRPNDIVDCLPLVHSAGNIVGICQLLHGMVTVIRKFSASRFWDDCIKYNCTIVQYI 334  
DB 247 GVHOEDVILALPLVHMSGSLIGVGLGIGATVVLKPKFSASFWDCCQKRVTVFQYI 306  
QY 335 GELCRYLLNQPRAENQHOVRMALGNLROSINWTFSSRPHIPQVAEFTYGATECNSLG 394  
DB 307 GELCRYLVNQPSPKAEFDHKVRLAVGSLRPDTWERFLRRFGPLQILETYGTEGNVATF 366  
QY 395 NFDQVAGCGFNSRLISFVPIRLVRVNEDTMELIRPDGVCIPCQGPGLVGRITQK 454  
DB 367 NYTGRQGVGRASWLYKHLFFSLPIRLVDVMTGEPINRAQGHCMTSPGEPGLLVAPVSOQ 426  
QY 455 DPLRRFDGYNQGNANK-KIARDVFKGDQAYLTGDLVMDLGLYLFVFRDRTGTPRWKG 513  
DB 427 SP---FLGYAGAPELAKLLKDVFWSGDVFNTGDLVCDQGFLEHFRDGTIRWKG 483  
QY 514 ENVSTVEGTLRLDMDADVGVGVPGTEGGRAGMAVA-SPTGNCDLREFAQVLEKE 572  
DB 484 ENWATEVAEVLLETDFLQEVNIYGVTVPGHEGGRAGMAALRPPQALNLVQLVSHVSEN 543  
QY 573 LPLYPARIFLRLPLHKTGYKFKQTELKKEGDPDAIVKDPFLFYLDAAQGRVVPDQEA 632  
DB 544 LPPYARPRFLRQESLATTETFKQKVRMANEGDPDPSLDPLYVLDDQIDGAYLPLTPAR 603  
QY 633 YSRIQAGEEKL 643  
DB 604 YSALLSGDLRI 614

RESULT 10  
O70550

ID O70550 PRELIMINARY; PRT; 620 AA.  
AC O70550;  
DT 01-AUG-1998 (TEMBLrel. 07, Created)  
DT 01-AUG-1998 (TEMBLrel. 07, Last sequence update)  
DT 01-OCT-2000 (TEMBLrel. 15, Last annotation update)  
DE VERY-LONG-CHAIN ACYL-COA SYNTHETASE.  
GN VLACS.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BALE/C; TISSUE=WHOLE LIVER;  
RX MEDLINE=98218572; PubMed=9559670;  
RA Berger J., Truppe C., Neumann H., Forss-Petter S.;  
RT "cDNA cloning and mRNA distribution of a mouse very long-chain acyl-  
CoA synthetase";  
RL FEBS Lett. 425:305-309(1998).  
DR EMBL; AJ223958; CAA11687.1; -  
DR InterPro: IPR000873; -  
DR Pfam: PF00501; AMP-binding; 1.  
DR PROSITE; PS00455; AMP\_BINDING; 1.  
SQ SEQUENCE 620 AA; 70394 MW; FC59270C41784364 CRC64;

Query Match 30.2%; Score 1021; DB 11; Length 620;  
Best Local Similarity 37.7%; Pred. No. 5.1e-77;  
Matches 235; Conservative 108; Mismatches 248; Indels 32; Gaps 12;

QY 32 LYLGGSGWFRIRVFTK-----IRRDFGLVLLKVKARQCLQER--RTVPILFASTVR 85  
DB 5 LYTGLAGLLLPULLTCCPYLLQDVYFLRLANMARVRSYRQRRPVTRITRAFLQAR 64  
QY 86 RHPDRTALFECTDHTWFRQLDEYSSVANFLOAR-GLASGDVAALFMENRNEFVGLWL 144  
DB 65 KTHPFPFLFR--DETLYAQVDRSNQVARALHDQLRGQDCVALFMGNEPAYVWIL 122  
QY 145 GMAKLGVEAALINTNLRRDALLHCLTTSARALVFGSEMAISCEVHASL--DPSLSLFC 202  
DB 123 GLKLGCPMACLINIRAKSLHCFQCCGAKVLLASPDQEAPEALPTLKKDASVFFV 182  
QY 203 SCSWEPGAVPSTEHLDPL-LKDPKHLPCDKGFTKLFYIYTSGTGLPKAAIYVHS 261  
DB 183 SPTSNTNGVDITLDKVDGVSAPETPSWRS--EVTFTPAVYIYTSGTGLPKRAATINH 240  
QY 262 RYRMAALVYGFRA-----RPNDIVDCLPLVHSAGNIVGICQLLHGMVTVIRKFS 315  
DB 241 R-----LWYGTGLAMSSGITAQDVIYTTMPLVHSAALMIGLHGCIVVGATLALRSKS 293  
QY 316 ASRFWDDCIKYNCTIVQYIGELCRYLLNQPRAENQHOVRMALGNLROSINWTFSSRF 375  
DB 294 ASQFWDCCRYNVTVIQYIGELLYLNTPOKPNDRDHKKVKKALGNLGDVWREFIKRF 353  
QY 376 HIPQVAEFTYGATECNSLGNFDSQVAGCFNSRLISFVPIRLVRVNEDTMELIRPDGV 435  
DB 354 GDHIVYEFTASTEGNIGFVNYPRKIGAVGRANVLRQVARELYKIDVDEKDEPVDANGY 413  
QY 436 CIPCQGPGLVGRITQKDLRPFDDGYL-NOGANNKKIAKDVFKGDQAYLTGDLVMD 494  
DB 414 CLKVPKGVGLVCKLTQLTPT---FIGAGGKTQTEKKRLRDFVFKKGDYIFNFGDGLIMD 470  
QY 495 ELGYLYFRDRTGDTFRKGENVSTVEGTLRLDMDADVGVGVPGTEGGRAGMAAV- 553  
DB 471 RENFVYFHDVGRDIFRKGENVATEVADIVGLVDFVEEVNYGVPPVPGHEGRIGMASLK 530  
QY 554 ASPTGNCDLERFAQVLEKELPLYPARIFLRLPLHKTGYKFKQTELKKEGDPDAIVKD 613  
DB 531 IKENYEFNGKKLFQHIAEYLPYARPRFLRIQDTIETGTGTFKKRKYTLMEEGFNPTVKD 590  
QY 614 PLFYLDAAQGRVVPDQEAYSRI 636  
DB 591 TLYFMDAETTFVPMENIYNAI 613

## RESULT 11

Q9Y2P5 PRELIMINARY; PRT; 690 AA.  
AC Q9Y2P5; 01-NOV-1999 (TREMELrel. 12, Created)  
DT 01-NOV-1999 (TREMELrel. 12, Last sequence update)  
DT 01-MAR-2001 (TREMELrel. 16, Last annotation update)  
DE VERY LONG-CHAIN ACYL-COA SYNTHETASE HOMOLOG 2.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Primates; Catarrhini; Hominoidea; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Steinberg S.J., Wang S.J., Watkins P.A.;  
RT "Human Very Long-Chain Acyl-CoA Synthetase Homolog 2";  
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF064255; RAD29444.1; -  
DR InterPro: IPR000873;  
DR Pfam: PF00501; AMP-binding; 1.  
DR PROSITE: PS00455; AMP-BINDING; 1.  
SQ SEQUENCE 690 AA; 75384 MW; 011313424D794546 CRC64;

Query Match 29.1%; Score 984.5; DB 4; Length 690;  
Best Local Similarity 37.7%; Pred. No. 7e-74; Indels 41; Gaps 15;  
Matches 252; Conservative 93; Mismatches 283;  
QY 2 LLGASLVGVLLSKVLK---PWTQVGSLLFLYLK-----SGGWRFRVFIKTIR 50  
DB 36 LLGPTCCVLLGLAMLPWLPWPHGLSLAAALATLLPARLPGLRWLPA----- 89  
QY 51 RDIPGGVLVLKAKVRQCL--QERRTPVILPASTVRRHPDKTALLFEGETDT-HWTFROL 107  
DB 90 -DVIFLAKILHLGLKIRGLSROPDPDFVDAFERRARAQPGALLVWTPGAGSVTFGEL 148  
QY 108 DEYSSSVANFLOAR-----GLAGDVAIEMENRNEFVGL--MLGMKLGVEAALINTNL 160  
DB 149 DARACQAAWALKAEGLDPASLCAGEPTALLVLAQVAPALCMWLGAKLGCPTAWINPHG 208  
QY 161 RRDALLHCLTTSRRALVFGSEMAICEVHASLD-PSLSLFCSSGSGWEPGAVPPSTEHL 219  
DB 209 RGMPLAHSLSSGARVLVWDPDLRESLEELPKLOAENIRCFYLSHTSP---TPGVGALG 265  
QY 220 PLLKDAKPH-LPSCPDKGFT--DKLEFYITSGTGLPKAAIVVHSRYRMAALVYIGERM 276  
DB 266 AALDAAPSHVPADLRAGITWRSPALFYITSGTGLPKALLHTRVLOMSKMLSL-GA 324  
QY 277 RPNDIVDCLPLXHSAGNIVGICQLLHGMVTVIRKFSASRFWDDCIKYNCTIVYIGE 336  
DB 325 TADDVYVTVPLXHVMLGVVGLGCLDLGATCVLAPKFTSCFWDCCRQHGVTVILYVGE 384  
QY 337 LCRYLLNPPREANQHVQWALNGRLQSIWTFNSSRFHIPQVAFYFGATECNSIGNF 396  
DB 385 LLRYLCNIPQPEDRTHVRLAMNGRLADVWETFOQRFPIRIWEVYSGTEGMMGLVNI 444  
QY 397 DSQVGAFCNSKILSFYPIRLVRVNEDTMELIRGPDGVCIPQPGEPGLVGRITQKDP 456  
DB 445 VRCGALGKMSCLRLMLSPFELVQFDMEAEPVEDNQGFCIPVGLGEPGLLTKVYSQQP 504  
QY 457 LRREDGYLN-QGANKKIADKVFKKGOAYITGDVLVMDLGLYIFRDRGTDFRWKNGEN 515  
DB 505 ---FVGYRGPRELSERKILVRNROSGDVYNTGDVLAMDREGLYIFRDLGDTFRWNGEN 561  
QY 516 VSTTEVEGTLRLDMDADVGYVEVPGTEGRAGMAAV-ASPTGNCDLERFAQVLEKELP 574  
DB 562 VSTHEVEGVLSDVFLQGVNNGVVCVPGCEGKVGMAAVQLAPGQTFDGEKLYQHVRWLP 621  
QY 575 LYARPIFLRLPELHKTGYTKFQKTELKKEGFPDPAIVKDPFLFYLDAGKGRVPLDQAYS 634  
DB 622 AYATPHFIRIQDAMEVSTFKMLKTRLVREGFNVGIVVDPLFLVLDNRAGSFRPLTAEMQ 681

QY 635 RIQAGEEKL 643  
DB 682 AVCEGTWRL 690

## RESULT 12

Q9ES38 PRELIMINARY; PRT; 690 AA.  
AC Q9ES38; 01-MAR-2001 (TREMELrel. 16, Created)  
DT 01-MAR-2001 (TREMELrel. 16, Last sequence update)  
DT 01-MAR-2001 (TREMELrel. 16, Last annotation update)  
DE BILE ACID COA LIGASE.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA STRAIN=SPRAGUE-DAWLEY;  
RC Falany C.N., Xie X., Wheeler J., Wang J., Barnes S.;  
RT "Molecular cloning and expression of rat liver bile acid CoA ligase";  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF242189; AAG09770.1; -  
KW Ligase.  
SQ SEQUENCE 690 AA; 76265 MW; D5BEB8317758DA59 CRC64;

Query Match 28.6%; Score 969.5; DB 11; Length 690;  
Best Local Similarity 37.8%; Pred. No. 1.3e-72; Indels 29; Gaps 12;  
Matches 233; Conservative 106; Mismatches 249;  
QY 46 IKTIRRDIFGVLVLLKVKAKVRQCLQERRTPVILFASTVRR---HPDKTALIFEGETD-T 100  
DB 84 LRLWHLKDVAFKLLFYGLNLRRL--NRHPELFVDALEQQAQARPDVALVCTGSEGC 141  
QY 101 HWTFQQLDEYSSSVANFLOARGLAS-----GDVAALFEMENR-ETVGLWLGMAKLGVEA 153  
DB 142 SITNELNAKACQAAWALKAEKATQEDKGATAILVPSKISALSLSVFLGAKLGPV 201  
QY 154 ALINLNRDALLHCLTTSRRALVFGSEMAICEVHAS-LDPSLSLFCSSGSGWEPGAVP 212  
DB 202 AWINPHSGMPLHSSVQSSGASVLVDPDLQENLEEVLPKLAENIRCFYLGHSPP---T 258  
QY 213 PSTEHLDPKLDAPKHLPSCPDK-----GFTDKLFYITSGTGLPKAAIVVHSRYRMA 267  
DB 259 PGVEALGAALDAAPS--DPVPAKLKRAIKKSPAFIYITSGTGLPKPAILSHERVIOIS 316  
QY 268 ALVYVGFMRPNDIVDCLPLXHSAGNIVGICQLLHGMVTVIRKFSASRFWDDCIKYN 327  
DB 317 NVLSFCGR-TADDVYVTVPLXHVMLGVVGLGCLDLGATCVLAPKFTSASRYWAEQYS 375  
QY 328 CTIVQYIGELCRYLLNQPPEAENQHVQWALNGRLQSIWTFNSSRPHIPQVAFYFGAT 387  
DB 376 VTIVVLYVGEVLYLCVNCVPGQPEDKKHTVRFALNGRLADVWENFQOQRFPIQIWEIYGT 435  
QY 388 EGNCSLGNFDSQVGAFCNSKILSFYPIRLVRVNEDTMELIRGPDGVCIPQPGEPGL 447  
DB 436 EGNVGLMNTVGHGAVGKTSCTCFIRMLTFLVQFDIETAEPPVRDKQKQFCIPVETGKPGLL 495  
QY 448 VGRITQKDLRRFDGCVLNGQANNNKIADKVFKKGOAYITGDVLVMDLGLYIFRDRGTGD 507  
DB 496 LTKIRKNQPFYLRG--SQDETFRKLVRANVQVGLDYNTGDVLALDQEGFFYFRDLGD 553  
QY 508 TFRWNGENVSTTEVEGTLRLDMDADVGYVEVPGTEGRAGMAAV-ASPTGNCDLERFA 566  
DB 554 TFRWNGENVSTTEVEGTLRLDMDADVGYVEVPGTEGRAGMAAV-ASPTGNCDLERFA 613  
QY 567 QVLEKELPFLYARPIFLRLPELHKTGYTKFQKTELKKEGFPDPAIVKDPFLFYLDAGKGRV 626  
DB 614 QHVRSLPYPATPHFIRIQDAMEVSTFKMLKTRLVREGFNVGIVVDPLFLVLDNRAGSFRPLTAEMQ 673





